ANALYSIS OF THE QUALITY OF IMAGE DATA ACQUIRED BY THE LANDSAT-4 THEMATIC MAPPER AND MULTISPECTRAL SCANNERS

Principal Investigator

Professor Robert N. Colwell

Remote Sensing Research Program
Space Sciences Laboratory
University of California
Berkeley 94720

Period of Performance

January 3, 1983 - August 13, 1984

Quarterly Status and Technical Progress Report #3
July 1, 1983 - September 30, 1983

NASA Contract #NAS5-27377
National Aeronautics and Space Administration
Goddard Space Flight Center
Greenbelt, MD

October 11, 1983

TABLE OF CONTENTS

1.0 TECHNICAL PROGRESS REPORT

Our research during this quarter was focused on 1) preparation of the PECORA VIII symposium paper, 2) geometric analysis of TM film products and 3) design of interpretation tests to determine optimum TM band combinations for renewable resource inventories.

1.1 PECORA VIII Symposium

Based on the invitation to present our TM image quality research results at the PECORA VIII Symposium, a paper was prepared which summarized our results to date. This paper was compiled from previous quarterly reports and drafts of the Early Results Symposium paper. The major topics in the paper included (1) the early significant results on the improved capability of TM Band 5 to discriminate specific crop types which are indistinguishable on TM Band 3 or 4; (2) geometric analysis of TM and MSS film products; and (3) interpretation of renewable resource features on TM and MSS film products. This paper will be published in the proceedings of the PECORA VIII symposium in January 1984.

1.2 Geometric Analysis of TM and MSS Film Products

The geometric quality of the TM and MSS film products are being evaluated by making selective photo measurements such as scale, linear and area determinations; and by measuring the coordinates of known features on both the film products and map products, and then relating these paired observations using a standard linear least squares regression approach. Specifically, the TM and MSS "Sacramento" scene (2/1/83) is gridded into nine equal area blocks. Within each block 50 to 60 points will be selected which are adequately represented on both film and map products. The points, or portions of natural features, will be measured in terms of relative image coordinates and corresponding map coordinates extracted from USGS 7 1/2" topographic quadrangle maps.

Half of these ground control points will be randomly selected to develop the regression relationship between the image and map coordinate systems. The geometric fidelity of the imagery will be based on the remaining "test" points and illustrated by a) plotting the magnitude and direction of the residual errors and b) performing an Analysis of Variance (ANOVA) on these residuals using each block of the image grid as a treatment. The range in elevation in the Sacramento scene is approximately 1000m., and the analysis by block will provide a useful assessment of the geometric quality of the full scene for a wide range of elevation conditions.

1.3 Interpretation of Resource Features on TM and MSS Film Products

To evaluate the quality and utility of the TM film products and various band combinations for detecting and identifying important forest and agricultural features, quantitative interpretation tests will be developed and conducted using skilled image analysts. The TM band combinations to be tested include, but are not limited to, Bands 1, 2, 3; Bands 2, 3, 4;

Bands 1, 3, 4; Bands 3, 4, 5; and Bands 2, 4, 5. For each band combination and for each major resource category, selective photo interpretation keys will be developed for use by the image analysts.

On each of the TM color composites, points will be located for each resource category. The points will be randomly located on each image type so that "interpreter recall" will not bias the test results. To administer the tests, a pool of 12 interpreters will be divided into three, four-person groups, and the tests administered during three 120-minute sessions. For each test, the interpreter will be asked to record the label or identifier code, for all points selected and annotated on the color composites. After testing is completed, the answer sheets will be analyzed and error matrices generated for each interpreter for all band combinations. The resulting error matrices will then be aggregated to produce an error matrix based on the total interpreter responses obtained for each band combination.

Percent correct and percent commission errors will be calculated for the error matrices as follows:

$$\text{Percent correct} = \frac{\text{number of correct interpretations of a resource category}}{\text{total number of that resource category present}} \times 100$$

$$\text{Percent commission error} = \frac{\text{number of incorrect interpretations of a resource category}}{\text{total number of that resource category indicated by the interpreter}} \times 100$$

In addition, for each matrix a Kappa statistic will be calculated as follows:

$$\text{Kappa} = \frac{\Sigma \text{Diagonal}/N - \Sigma(\text{row total} \times \text{column total})/N^2}{1 - \Sigma(\text{row total} \times \text{column total})/N^2}$$

where: N = total number of observations in the matrix

The Kappa statistic, which is a non-parametric measure of agreement between "ground truth" and photo interpreter labels, is used to rank the error matrices. The rankings are considered to be significantly different based on the values obtained through use of the following relationship:

$$\text{Delta Kappa} = \frac{|\text{Kappa}_i - \text{Kappa}_j|}{(\text{Variance}_{\text{Kappa}_i} + \text{Variance}_{\text{Kappa}_j})^{1/2}}$$

If the calculated Delta Kappa between two matrices exceeds 1.96, we conclude that Kappa values are significantly different at the 95 percent confidence level; if the calculated value is less than 1.96, we conclude that the Kappa values are not significantly different and that the band combinations represented by those matrices were equally interpretable. A complete description of the use of the Kappa statistic can be found in Congalton, et al., 1981.

It should be emphasized that the Kappa statistic cannot be used alone when evaluating the usefulness of different image combinations. While it is a very helpful statistical tool for the ranking combination, because it combines errors of omission and commission, it only measures the degree of association along the diagonal of the error matrix. It is very possible that other associations may exist in this matrix. For example, an interpreter may consistently misidentify the mixed conifer category as the true fir category, and vice versa. While such an error would result in a very low Kappa value for the image type, it is really a reflection on interpreter skill and not on image quality. In this case the interpreter, with the available training aids and background experience, has confused the identifying characteristics of one conifer type with another. Since he has done it consistently, it is merely a matter of retraining him. Therefore, it is very important to not only look at the Kappa statistic, but also at the percent correct value and the percent omission-error and commission-error values, as given by the individual entries in the error matrix, in order to properly evaluate an image type for a given application (Colwell, et al. 1982).

1.4 References

Colwell, R. N., A. S. Benson, K. J. Dummer, and J. T. Hardin. 1982. Analysis of Wildland Data Acquired by Advanced High Altitude Systems on the San Juan National Forest, Colorado. Final Report, U.S. Forest Service Contract #53-3187-1-40.

Congalton, R. G., R. A. Mead, R. G. Odewald, and J. Heinen. 1981. Analysis of Forest Classification Accuracy. Cooperative Research Report. U.S. Forest Service, Virginia Polytechnic Institute and State University, Blacksburg, VA.

2.0 PUBLICATIONS AND PRESENTATIONS

There were no publications or presentations during this three-month period of performance.

3.0 FUNDS EXPENDED TO DATE

The funds expended to August 31, 1983 under this contract are summarized in NASA Form 533M, "Monthly Contractors Financial Management Report", dated 10/7/83.

4.0 PROBLEMS ENCOUNTERED TO DATE

Specific problem areas for this reporting period are discussed in the July, August, and September monthly reports submitted to Mr. Darrel Williams, Code 923, NASA-Goddard Space Flight Center.

DISTRIBUTION

10 - TECHOFF, Code 902, NASA-GSFC

1 - CGOFF, Code 984.4, NASA-GSFC

1 - Publications Section, Code 253.1, NASA-GSFC

1 - Patent Counsel, Code 204, NASA-GSFC

1 - Contracts and Grants, SSL: U.C. Berkeley

5 - RSRP, SSL: U.C. Berkeley